

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2004, 11:20:12 ; Search time 279.4 Seconds

(without alignments)
2635.552 Million cell updates/sec

Title: US-10-062-875A-5

Perfect score: 18
Sequence: 1 ggcgataatcaacgggtga 18

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl: *
1: gb_ba: *
2: gb_htg: *
3: gb_in: *
4: gb_cm: *
5: gb_ov: *
6: gb_pat: *
7: gb_ph: *
8: gb_pl: *
9: gb_pr: *
10: gb_ro: *
11: gb_sts: *
12: gb_sy: *
13: gb_un: *
14: gb_vi: *
15: em_ba: *
16: em_fun: *
17: em_hum: *
18: em_in: *
19: em_mu: *
20: em_om: *
21: em_or: *
22: em_ov: *
23: em_pat: *
24: em_ph: *
25: em_pl: *
26: em_ro: *
27: em_sts: *
28: em_un: *
29: em_vi: *
30: em_htg_hum: *
31: em_htg_inv: *
32: em_htg_other: *
33: em_htg_mus: *
34: em_htg_pln: *
35: em_htg_rod: *
36: em_htg_mam: *
37: em_htg_vrt: *
38: em_sy: *
39: em_htgo_hum: *
40: em_htgo_mus: *
41: em_htgo_other: *

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	18	100.0	1306	1	BPIS1001	X66858 B.parapertu
2	16.4	91.1	12598	1	AE014723	AE014723 Bifidobac
3	16.4	91.1	129302	8	AC133778	AC133778 Oryza sat
4	16.4	91.1	136321	8	AP003300	AP003300 Oryza sat
5	16.4	91.1	138701	8	AP003243	AP003243 Oryza sat
6	16.4	91.1	163795	9	CNS05TEU	AL359396 Human chr
7	16.4	91.1	164236	8	AC133339	AC133339 Oryza sat
8	16.4	91.1	168896	2	OSJN00285	AL731640 Oryza sat
9	16.4	91.1	300115	1	AP005076	AP005076 Vibrio pa
10	16.4	91.1	349980	6	AX492784	AX492784 Sequence
11	16.4	91.1	349980	6	AX492785	AX492785 Sequence
12	16.4	91.1	349980	6	AX553951	AX553951 Sequence
13	16.4	91.1	349980	6	AX553952	AX553952 Sequence
14	15.4	85.6	2814	3	AF400665	AF400665 Caenorhab
15	15.4	85.6	2814	6	AX351522	AX351522 Sequence
16	15.4	85.6	55169	3	AC006605	AC006605 Caenorhab
17	15.4	85.6	76396	2	AC012995	AC012995 Drosophila
18	15.4	85.6	90689	2	AC141010	AC141010 Rattus no
19	15.4	85.6	94225	2	AC138542	AC138542 Magnapor
20	15.4	85.6	153205	2	AP004552	AP004552 Oryza sat
21	15.4	85.6	154881	2	AP005919	AP005919 Oryza sat
22	15.4	85.6	190801	3	AC009212	AC009212 Drosophila
23	15.4	85.6	294050	1	AL627268	AL627268 Salmonella
24	15.4	85.6	301983	1	AE016840	AE016840 Salmonella
25	15.4	85.6	303432	3	AE003604	AE003604 Drosophila
26	15.4	85.6	322194	8	CNS09S4S	EX088700 DNA cent
27	15	83.3	1474	6	I65498	I65498 Sequence 8
28	15	83.3	1474	6	I90341	I90341 Sequence 6
29	15	83.3	1474	8	CHU39834	U39834 Cuphea hook
30	15	83.3	2053	14	AF235166	AF235166 American
31	15	83.3	20366	1	AE008885	AE008885 Salmonella
32	15	83.3	96086	1	STR5TMD1	AF233324 Salmonella
33	15	83.3	125646	2	AC137986	AC137986 Medicago
34	15	83.3	169547	10	AL590503	AL590503 Mouse DNA
35	15	83.3	258050	1	AL627278	AL627278 Salmonella
36	15	83.3	299991	1	AE016845	AE016845 Salmonella
37	14.8	82.2	400	11	G16871	G16871 human STS. S
38	14.8	82.2	690	6	AX392266	AX392266 Sequence
39	14.8	82.2	952	8	AY261522	AY261522 Suaeda ma
40	14.8	82.2	1233	1	AF273254	AF273254 Coxsiella
41	14.8	82.2	1481	4	AY240020	AY240020 Bos tauru
42	14.8	82.2	1659	6	AX189037	AX189037 Sequence
43	14.8	82.2	1967	1	ECSRMB	X14152 E. coli Strm
44	14.8	82.2	2000	6	AX508239	AX508239 Sequence
45	14.8	82.2	2186	8	AF024625	AF024625 Brassica

ALIGNMENTS

RESULT 1
BPIS1001
LOCUS
DEFINITION B.parapertussis insertion sequence IS1001 tnpA gene for
ACCESSION X66858 SS1601
VERSION X66858.1 GI.39755
KEYWORDS insertion sequence; insertion sequence IS1001; tnpA gene;
transposase.
SOURCE
ORGANISM Bordetella parapertussis
Bordetella parapertussis
Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
Alcaligenaceae; Bordetella.
REFERENCE
AUTHORS 1
van der Zee, A., Agterberg, C., van Agterveld, M., Peeters, M. and

TITLE Mooi,F.R.
Characterization of ISI001, an insertion sequence element of
Bordetella parapertussis
JOURNAL J. Bacteriol. 175 (1), 141-147 (1993)
MEDLINE 93106948
PUBMED 8093238
REFERENCE 2 (bases 1 to 1306)
AUTHORS van der Zee,A.
TITLE Direct Submission
JOURNAL Submitted (12-JUN-1992) A. Van Der Zee, National Institute of Health
&, Environmental Protection, A Van Leeuwenhoeklaan 9, P O Box 1,
3720 BA Bilthoven, THE NETHERLANDS
COMMENT On Jul 10, 2002 this sequence version replaced gi:262329.
FEATURES
source
1..1306
/organism="Bordetella parapertussis"
/mol_type="genomic DNA"
/db_xref="taxon:519"
/clone="pRPP1"
1..1306
/insertion_seq="ISI001"
83..1303
/gene="TnpA"
83..1303
/gene="TnpA"
/codon_start=1
/transl_table=11
/product="transposase"
/protein_id="CAA47326.1"
/db_xref="GI:39756"
/db_xref="SWISS-PROT:Q06126"
/translation="MLDRKLMSIGWGQYGVREVMPEDPGRITLSYLPKPAKWL
ECCGACRQVHETVYRVYRDLPFIEYRVVLHVPRRLWCQCGGPRLERLAMLGRIYQ
VTDRLAQACSQLGSSNVQAVARFFELGWTVTKLIDKARLRSVREPDMSKIEYLMD
EFALYKHGRYATVYVDPIGRQVLMIGPGRSRETARAFPEQLPPGAQRKIKAVAI
AYLEIQAHSPQAEIYVDLHVAVAKYGREVIDRVRDQANOLRQDRPARRIKSRML
LIRNRDNLDRQAVRLDELLOANQPLTYVLRDELKRLMTYQRPAMARQAMNHWYEQ
AEOSGIAALNTFAORLKGYLHGLILARCRHPLNTSIVEGINNTIKVIXRAYGYRQY
FFLKIRAAFPGNAR"
BASE COUNT 269 a 371 c 420 g 246 t
ORIGIN
Query Match 100.0%; Score 18; DB 1; Length 1306;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCGGATATCAACGGGTGA 18
Db 375 GCGGATATCAACGGGTGA 392
RESULT 2
AE014723/c 12598 bp DNA linear BCT 31-OCT-2002
LOCUS Bifidobacterium longum NCC2705 section 110 of 202 of the complete
DEFINITION genome.
ACCESSION AE014723 AE014295
VERSION AE014723.1 GI:23326232
KEYWORDS
SOURCE Bifidobacterium longum NCC2705
ORGANISM Bifidobacterium longum NCC2705
Bacteria; Actinobacteria; Actinobacteridae; Bifidobacteriales;
Bifidobacteriaceae; Bifidobacterium.
REFERENCE 1 (bases 1 to 12598)
AUTHORS Schell,M.A., Karmirantzou,M., Snel,B., Vilanova,D., Berger,B.,
Pessi,G., Zwahlen,M.-C., Desiere,F., Bork,P., Delley,M.,
Pridmore,D. and Arigoni,F.
TITLE The genome sequence of Bifidobacterium longum reflects its
adaptation to the human gastrointestinal tract
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (22), 14422-14427 (2002)
PUBMED 12381787
REFERENCE 2 (bases 1 to 12598)
AUTHORS Schell,M.A., Karmirantzou,M., Snel,B., Vilanova,D., Berger,B.,

TITLE Pessi,G., Zwahlen,M.-C., Desiere,F., Bork,P., Delley,M.,
Pridmore,D. and Arigoni,F.
JOURNAL Submitted (27-AUG-2002) Bioscience, Nestle Research Center, P.O.
Box 44, Lausanne 26 1000, Switzerland
FEATURES
source
1..12598
/organism="Bifidobacterium longum NCC2705"
/mol_type="genomic DNA"
/strain="NCC2705"
/db_xref="taxon:206672"
/complement(85..678)
/gene="BL0998"
complement(85..678)
/gene="BL0998"
/note="COG family: transcriptional regulators, similar to
M. xanthus card; PFM_ID: TF_Card"
/codon_start=1
/transl_table=11
/product="possible transcriptional regulator"
/protein_id="AA24806.1"
/db_xref="GI:23326233"
/translation="MSYKVGDMVYPRHGAQVEATERVYGVTRREYLQSLVSSDG
LVINVPVDNAKRVGRDIVSASEVAKVFDILRTPIIEKMWMSRYKLVNVEKIATGV
NKIAEVVRDLAQRDVDEHGLSAGEKMLTKARAILTSEIALSEKIDBAEAOQLDVLN
GYEPAQPGDDKHTAEPBEAADTLARVEAEKSKK"
895..3108
/gene="g1gB"
/note="synonym: BL0999"
895..3108
/gene="g1gB"
/note="COG family: 1,4-alpha-glucan branching
enzyme; PFM_ID: isoamylase_N; PFM_ID: alpha-amylase"
/codon_start=1
/transl_table=11
/product="1,4-alpha-glucan branching enzyme"
/protein_id="AA24807.1"
/db_xref="GI:23326234"
/translation="MPVAQGLDLSNAEFYNPHGVLGHLGKHAATATIRVLRPL
AKSVTILLTODGEYPMTHENYNGVFVTVASGCKKOPTIPDYRVTEWESGAVLIEDP
YRMPPTVGDMDTYLFEGGRHEKLMETLGAHVLRYPDPMGAGDTGPEQVGTAFSVA
PNAHAVRVVGNFNAMDGRRHAMRELSSGVWEIFIPGAGETVYKQILNANTWEMK
ADPMERQHEVPPSTASIVTESYTKWMDMAWQHRRITNPHDGPVSIYEVHAGSKQGL
TYRDLAKQLVDYKQEGFTHEVEMPLAQHPFSGSWGYQVGYAVDSRLGSPDDFRYL
VDQFHQAGIVMDVWPVPAHFPKDAFAGRFDTPLYEDPDPLRGEHPWGTYVFNGR
REVRFNLVANAFLMLIEDLHVDAALRVAVSSMLYLDYSREPGHMRPNIVYGRNLEAID
FLKEATATAYAKNPGVWVIAESTAWPGITAPTSAGGIFGGMKMMGMHDTLEYLHE
EPINRKWHNEITPSMYYAYSEHYVLPISHDEVYVGKSLYKMGDDDMQKLAVGRSM
FAYQWAHPGKXLSFNGNELAQWGEWDHDSIDWDCLNMQEHROQTWADLNAFYKAH
PALMSQDFDPAGFQWLTSDDADHNTLSFLRIGTKGETLAVTVNPSGEAMSDYQVALPT
GGKWTVEFTTDDAKYGGSDIHNGTFEAVEGEYHSRPFSAKITVPALGVFLKPED"
3144..3866
/gene="BL1000"
3144..3866
/gene="BL1000"
/note="COG family: response regulators consisting of a
Chey-1-like receiver domain and a HTH DNA-binding domain;
PFM_ID: response_reg; PFM_ID: trans_reg_C"
/codon_start=1
/transl_table=11
/product="response regulator of two-component system"
/protein_id="AA24808.1"
/db_xref="GI:23326235"
/translation="WLNNTARQNTTPRVLVVEDEPTLATAIQRIATGWTARVAGD
GASAVQAASQRLPDLVIMDMLPVMDGLAETKRIVARFPVALLTARDDADKVGIL
GAGADVMTKPFSEMBELIARCKALLRVRERAKVIAKSENEXLDFGSMVIDPAQRIV
TVNGEQVHLTPTEFDLLATLARPKSVLTREKLLBEVWDVWVDSGTRTVDSHVKALRH
KLGADTIRTVHGVGAFEPPTA"
3863..4939
/gene="BL1001"
3863..4939
/gene="BL1001"
/note="COG family: sensory transduction histidine kinases;

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2004, 11:20:12 ; Search time 294.922 Seconds

(without alignments)
2635.552 Million cell updates/sec

Title: US-10-062-875A-6

Perfect score: 19
Sequence: 1 cagggaactcgtccatc 19Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 segs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	19	100.0	1306	1 BPIS1001	X66858 B.parapertu
C 2	17.4	91.6	7021	1 AF019635	AF019635 Pseudomon
C 3	17.4	91.6	7109	1 PP041301	U41301 Pseudomonas
C 4	17.4	91.6	14861	1 AF043544	AF043544 Pseudomon
C 5	17.4	91.6	180699	9 AC079403	AC079403 Homo sapi
C 6	17	89.5	16644	1 AE006918	AE006918 Mycobacte
C 7	17	89.5	20760	1 MTCY21D4	Z80775 Mycobacteri
C 8	17	89.5	38721	1 MSGY219	AD000013 Mycobacte
C 9	17	89.5	138229	2 HSAC001236	AC001236 Homo sapi
C 10	17	89.5	343050	1 BX248334	BX248334 Mycobacte
C 11	16.4	86.3	588	11 G79512	G79512 S208P6023RE
C 12	16.4	86.3	2596	4 AF411535	AF411535 Microcyt
C 13	16.4	86.3	10333	1 AE006873	AE006873 Sulfolobu
C 14	16.4	86.3	12421	1 AE000943	AE000943 Archaeogl
C 15	16.4	86.3	70360	2 AC101122	AC101122 Mus muscu
C 16	16.4	86.3	70571	2 AC145071	AC145071 Mus muscu
C 17	16.4	86.3	99761	8 AC006161	AC006161 Arabidops
C 18	16.4	86.3	101284	8 ATAC009991	AC009991 Arabidops
C 19	16.4	86.3	103517	8 AC073395	AC073395 Arabidops
C 20	16.4	86.3	105795	1 SYCCPNC	D64001 Synchocyst
C 21	16.4	86.3	120648	8 AP003859	AP003859 Oryza sat
C 22	16.4	86.3	169284	5 AL953899	AL953899 zebrafish
C 23	16.4	86.3	216491	2 AC128110	AC128110 Rattus no
C 24	16.4	86.3	231157	2 AC111347	AC111347 Rattus no
C 25	16.4	86.3	254989	2 AC107773	AC107773 Mus muscu
C 26	16.4	86.3	259783	2 AC111985	AC111985 Rattus no
C 27	16	84.2	321	6 BD077656	BD077656 5'EST of
C 28	16	84.2	640	6 BD082462	BD082462 87 human
C 29	16	84.2	855	6 E14996	E14996 Synchococc
C 30	16	84.2	1341	9 BC012415	BC012415 Homo sapi
C 31	16	84.2	1358	9 HSY12478	Y12478 H.sapiens m
C 32	16	84.2	1374	6 BD082403	BD082403 87 human
C 33	16	84.2	1529	6 BD082463	BD082463 87 human
C 34	16	84.2	2491	10 BC031769	BC031769 Mus muscu
C 35	16	84.2	4943	6 E14999	E14999 Synchococc
C 36	16	84.2	8481	1 AB010691	AB010691 Synchococ
C 37	16	84.2	74355	1 AY120853	AY120853 Synchococ
C 38	16	84.2	135578	9 AF064861	AF064861 Homo sapi
C 39	16	84.2	153679	2 EX000987	EX000987 Danio rer
C 40	16	84.2	154523	2 BX005217	BX005217 Danio rer
C 41	16	84.2	301442	1 AE016798	AE016798 Vibrio vu
C 42	16	84.2	340000	9 HS21C079	AL163279 Homo sapi
C 43	15.8	83.2	309	1 RHMONDMA	M65246 Rhizobium 1
C 44	15.8	83.2	481	11 G97381	G97381 S209P6007RB
C 45	15.8	83.2	658	11 BV076785	BV076785 S212P6527

ALIGNMENTS

RESULT 1
BPIS1001/c 1306 bp DNA linear BCT 07-JUL-2002
LOCUS B.parapertussis insertion sequence IS1001 tnpA gene for
DEFINITION B.parapertussis insertion sequence IS1001 tnpA gene for
transposase.
ACCESSION X66858 S51601
VERSION X66858.1 GI:39755
KEYWORDS insertion sequence; insertion sequence IS1001; tnpA gene;
transposase.
SOURCE Bordetella parapertussis
ORGANISM Bordetella parapertussis
Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
Alcaligenaceae; Bordetella.
REFERENCE 1
AUTHORS van der Zee,A., Agterberg,C., van Agterveld,M., Peeters,M. and

TITLE Mooi, F.R.
Characterization of IS1001, an insertion sequence element of
Bordetella parapertussis
JOURNAL J. Bacteriol. 175 (1), 141-147 (1993)
MEDLINE 93106948
PUBMED 8093238
REFERENCE 2 (bases 1 to 1306)
AUTHORS van der Zee, A.
TITLE Direct Submission
JOURNAL Submitted (12-JUN-1992) A. Van Der Zee, National Institute of Health
& Environmental Protection, A Van Leeuwenhoeklaan 9, P O Box 1,
3720 BA Bilthoven, THE NETHERLANDS
COMMENT On Jul 10, 2002 this sequence version replaced gi:262329.
FEATURES
source
1. 1306
/organism="Bordetella parapertussis"
/mol_type="genomic DNA"
/db_xref="taxon:519"
/clone="pRPP1"
1. 1306
/insertion_seq="IS1001"
83. 1303
/gene="TnpA"
83. 1303
/gene="TnpA"
/codon_start=1
/transl_table=11
/product="transposase"
/protein_id="CAA47326.1"
/db_xref="GI:39756"
/db_xref="SWISS-PROT:Q06126"
/translation="MLDRKLMSGLGWMQGYGVERVWEPEDPGRTLSIYIKPTAKVMLC
EOCGARCRQVHETTVRRVRLPIFEYRVLVHPRRLWCEQCGPRLESLAMLGROYR
VTDLAQAQSOLQSSNVQAVRPFELGMVTKLIDKARLRASVREPDSKI EYLAMD
EFLAHKGRYATVVVDPIGRQVLMIGPGRSRETARAFPEQLPFGAQRITKAVALDMT
AYELIOASHPOAEIVYDLFHVYAKYGREVIDRVRVQANQLRQDRPAPRIIKSSRWL
LLRNDRNLDRQAVRLDELLOANDPLTIVVLRDELKLMFYQRPAMARQAMNHWEQ
AEQSGIAALNTPAQRLKGYLHGLARCRHPLNTSIVEGINNTIKVIKRAYGYRDOEY
PFLKIRAPFGNAR"
BASE COUNT 269 a 371 c 420 g 246 t
ORIGIN
Query Match 100.0%; Score 19; DB 1; Length 1306;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CAGGGCAAACCTGTCATC 19
|||||
Db 574 CAGGGCAAACCTGTCATC 556
RESULT 2 7021 bp DNA linear BCT 21-SEP-1997
AF019635/c Pseudomonas putida TOL plasmid pDK1 xy1U, benzyl alcohol
LOCUS dehydrogenase (xy1W), benzaldehyde dehydrogenase (xy1C), xylene
DEFINITION monooxygenase hydroxylase component (xy1M), and xylene
monooxygenase electron transfer component (xy1A) genes, complete
cds.
ACCESSION AF019635
VERSION AF019635.1 GI:2425076
KEYWORDS Pseudomonas putida
SOURCE Pseudomonas putida
ORGANISM Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
Pseudomonadaceae; Pseudomonas.
REFERENCE 1 (bases 1 to 7021)
AUTHORS Gilbert, M.M., McCormick, T. and Benjamin, R.C.
TITLE Pseudomonas putida TOL plasmid pDK1 O/P region and xy1Wcma
sequence
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 7021)
AUTHORS Gilbert, M.M., McCormick, T. and Benjamin, R.C.

TITLE Direct Submission
JOURNAL Submitted (15-AUG-1997) Biology, University of North Texas, Denton,
TX 76203, USA
FEATURES
source
1. 7021
Location/Qualifiers
/organism="Pseudomonas putida"
/mol_type="genomic DNA"
/db_xref="taxon:303"
/plasmid="TOL plasmid pDK1"
1612. 2010
/gene="xy1U"
1612. 2010
/gene="xy1U"
/function="unknown"
/codon_start=1
/transl_table=11
/protein_id="AAB70822.1"
/db_xref="GI:2425077"
/translation="MLYFLNSAQIGCFRHYVPOSGPAVAVTLGVCDAGLELSEGEPKQ
ASPTNCEBPXVIGVRAITRTICVAVIGFISRNKRVIDASPTVRLHQSSISLSTDYM
PRSATVYPACRAFARALVYRSHSCSDITN"
2132. 3178
/gene="xy1W"
2132. 3178
/gene="xy1W"
/function="benzyl alcohol + NAD+ = benzaldehyde + NADH +
H+"
/codon_start=1
/transl_table=11
/product="benzyl alcohol dehydrogenase"
/protein_id="AAB70823.1"
/db_xref="GI:2425078"
/translation="MLAIRKTAAGPGGLSVDDVADPKAHSSTVLVKVEAAGTGTDL
LTKWGEFAKRMKLPITLGHVSGVIEQVSDVKGLRPGMRVSLESHPGCTCYCR
GWAHVCPNTRYPGVDFDGPASFVVVPESVCWVPVPGIIPLOAMEPFGLAVHASLE
GSGVSGNLTVSGCCPIGLMNI AAKALGASKVIATDHPRLTAAKMGADCEINAT
ESVYKTVRSGLGERGVDAVDIYSGQAAALKEILQSTTHGGEIRLMGVPSHDIILNLE
EWLFKGLVRLGHRLLFETWERTSLTIVTGRVDSLIVSHRLPLSAEDAPFEMALNG
QSLKILFEPNGTCI"
3210. 4673
/gene="xy1C"
3210. 4673
/gene="xy1C"
/function="benzaldehyde + NAD+ = benzoic acid + NADH + H+"
/codon_start=1
/transl_table=11
/product="benzaldehyde dehydrogenase"
/protein_id="AAB70824.1"
/db_xref="GI:2425079"
/translation="MRETKQPIWYKVFSSNNWVEARGVAVVDPNSGDIIGITVA
NGEDVDAVNAKRAQKEWAI PFSERAIVRAAEKLKERHEFADWNVREGCAIRP
KGLWEAGIAYEQHQAGLAYLSNGTIFPSAVPGRMNLQRPVGVGVVIAIPMNFPLF
LAMRSVAPALALGNVAILKPDLOTAVTGALITAEI FSDAGMDGVLVLPFGADVGEA
MVANSGINMISFTGSTQVGRLLIGKCGMLKKVALELGNVYHVLVPDADLDGAVGSA
AMGTFLHOGQVCMAGRAHVLHRDVAQOYAEKLARAKNLVVGDPNSDRVHLGLINDK
QVVRVHALVESAGSAGAKVLAGGTYYQATVIMDVKPEMEYFKESEIFGPVAPITV
FDSIEAEIELANCSEYGLAASITRALATGLDI AKRLNTGWHINDQPINCEPHVPFG
GMGASGSGRFGPASIIEFTQSQWISWEKPNYPF"
4701. 5816
/gene="xy1M"
4701. 5816
/gene="xy1M"
/function="toluene + O2 + NADH + H+ = benzyl alcohol +
NADH + H2O"
/codon_start=1
/transl_table=11
/product="xylene monooxygenase hydroxylase component"
/protein_id="AAB70825.1"
/db_xref="GI:2425080"
/translation="MDTLRYLIPVVTACGLIGFYGYGWLGAATFPLMVLVDVIL
PKDFSARKVSPFPADLTQYQLPLMIGLYGLLVGVENGRIELSEPLQYAGCILSLAW
LSGVPTLPVSHMLHRHMLPRKMAQLAMFYGDPNRDIAHVNTHLYLDTPLDSDTF
YRGQTIYSFVISATVGSYDAIKIEAETLRKQSPWNLSNKTQYVALLALPGLVA

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2004, 11:20:12 ; Search time 372.533 Seconds
(without alignments)
2635.552 Million cell updates/sec

Title: US-10-062-875A-7
Perfect score: 24
Sequence: 1 gtcttcgaactgggttgcatatc 24

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues
Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenBml:
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rsd:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	24	100.0	1306	1 BPIS1001	X66858 B.parapertu
2	20.8	86.7	73800	1 AF286216	AF286216 Pseudomon
3	19.8	82.5	5307	6 AX276469	AX276469 Sequence
4	19.8	82.5	311143	1 AE016762	AE016762 Escherich
5	18.4	76.7	17277	1 AF237701	AF237701 Pseudomon
6	18.4	76.7	161005	2 AC116595	AC116595 Mus muscu
7	18.4	76.7	172114	10 AC121903	AC121903 Mus muscu
8	18.2	75.8	1042	1 AY228241	AY228241 Pseudomon
9	18.2	75.8	1305	1 AF285635	AF285635 Burkholde
10	18.2	75.8	4620	6 AX078429	AX078429 Sequence
11	18.2	75.8	34244	1 AOPCZA363	AJ223399 Amycolato
12	18.2	75.8	40430	1 AB107223	AB107223 Pseudomon
13	18.2	75.8	85237	9 HSDJ653C5	AL049743 Human DNA
14	18.2	75.8	93829	3 AC007122	AC007122 Drosophil
15	18.2	75.8	119163	9 AC119679	AC119679 Homo sapi
16	18.2	75.8	121414	2 AC017234	AC017234 Drosophil
17	18.2	75.8	173165	3 AC099005	AC099005 Drosophil
18	18.2	75.8	174673	9 AC019187	AC019187 Homo sapi
19	18.2	75.8	174822	9 AC013275	AC013275 Homo sapi
20	18.2	75.8	182560	3 AC099024	AC099024 Drosophil
21	18.2	75.8	189023	9 AL592114	AL592114 Human DNA
22	18.2	75.8	191240	2 AC027738	AC027738 Homo sapi
23	18.2	75.8	204207	2 AC023147	AC023147 Homo sapi
24	18.2	75.8	217586	2 AC016076	AC016076 Homo sapi
25	18.2	75.8	262533	3 AE003809	AE003809 Drosophil
26	18.2	75.8	295883	2 AC138459	AC138459 Mus muscu
27	18.2	75.8	304517	1 AE016789	AE016789 Pseudomon
28	18.2	75.8	310029	1 AE016866	AE016866 Pseudomon
29	18.2	75.8	310678	1 AE016873	AE016873 Pseudomon
30	18.2	75.8	328863	2 AL627202	AL627202 Homo sapi
31	17.8	74.2	81319	9 AC068756	AC068756 Homo sapi
32	17.8	74.2	150114	2 AC074004	AC074004 Homo sapi
33	17.8	74.2	156313	2 AC141774	AC141774 Apis mell
34	17.8	74.2	169892	9 AC068397	AC068397 Homo sapi
35	17.8	74.2	184489	2 AC027771	AC027771 Homo sapi
36	17.8	74.2	189622	9 AC116442	AC116442 Homo sapi
37	17.8	74.2	211441	2 AC130659	AC130659 Mus muscu
38	17.8	74.2	214166	9 AC009660	AC009660 Homo sapi
39	17.8	74.2	215084	2 BX248502	BX248502 Danio rer
40	17.8	74.2	219086	9 AC010147	AC010147 Homo sapi
41	17.8	74.2	268063	2 AC110912	AC110912 Mus muscu
42	17.8	74.2	272761	2 AC094657	AC094657 Rattus no
43	17.8	74.2	313450	1 AL596170	AL596170 Listeria
44	17.8	74.2	319630	6 AX413016	AX413016 Sequence
45	17.8	74.2	349980	6 AX417046	AX417046 Sequence

ALIGNMENTS

RESULT 1
BPIS1001
LOCUS BPIS1001 1306 bp DNA linear BCT 07-JUL-2002
DEFINITION B.parapertussis insertion sequence IS1001 tmpr gene for
transposase.
ACCESSION X66858.1 GI:39755
VERSION X66858.1 GI:39755
KEYWORDS insertion sequence; insertion sequence IS1001; tmpr gene;
transposase.
SOURCE Bordetella parapertussis
ORGANISM Bordetella parapertussis
Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
Alcaligenaceae; Bordetella.
REFERENCE 1
AUTHORS van der Zee, A., Agterberg, C., van Agterveld, M., Peeters, M. and

Moqi,F.R.
TITLE Characterization of IS1001, an insertion sequence element of
Bordetella parapertussis
JOURNAL J. Bacteriol. 175 (1), 141-147 (1993)
MEDLINE 93106948
PUBMED 8093238
REFERENCE 2 (bases 1 to 1306)
AUTHORS van der Zee,A.
TITLE Direct Submission
JOURNAL Submitted (12-JUN-1992) A. Van Der Zee, National Institute of Health
&, Environmental Protection, A Van Leeuwenhoeklaan 9, P O Box 1,
3720 BA Bilthoven, THE NETHERLANDS
COMMENT On Jul 10, 2002 this sequence version replaced gi:262329.
FEATURES
source
1..1306
/organism="Bordetella parapertussis"
/mol_type="genomic DNA"
/db_xref="taxon:519"
/clone="prp1"
1..1306
/insertion_seq="IS1001"
83..1303
/gene="TnpA"
83..1303
/gene="TnpA"
/codon_start=1
/transl_table=11
/product="transposase"
/protein_id="CAA47326.1"
/db_xref="GI:39756"
/db_xref="SWISS-PROT:Q06126"
/translation="MLDRKIMESLGWGYGVERVWEPEDPGRITSLYKPTAKVMLC
EQGACRCQVETTVRVRDLPIFEYRVVLHPRRLWCEQCGPRLERLAMLGRYOR
VTDRLAQACSQLQSSNVOAVARFELGWHVTKTLDKARLASVREPWSKIEVLA
EFALHKGHRVAVVVDPIGRQVLMIGBGRSRETARAFEPQLPPGAQRIRKAVLDMTT
AYELEIQASHPOAEIVYDLFHVVAKYGREVIDRVRVDOANQLRODRPARRIKSSRWL
LLNRNDNLDRQAVRLDELLOANQLITVYVLRDELKRLMFYORPAMARQAMNHWEY
AEQSGIALNTFAORLKGYLHGLIARCRHPLNTSIVEGINNTIKVIRKRAYGYRDQY
FFLKIRAAFPGNAR"

BASE COUNT 269 a 371 c 420 g 246 t
ORIGIN

Query Match 100.0%; Score 24; DB 1; Length 1306;
Best Local Similarity 100.0%; Pred. No. 0.13;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTTCTTGAAGTGGGTGGCATAC 24
|||
Db 457 GTTCTTGAAGTGGGTGGCATAC 480

RESULT 2 73800 bp DNA linear BCT 24-MAR-2003
AF286216
LOCUS
DEFINITION Pseudomonas syringae pv. syringae syringopeptin synthetase (sypA),
syringopeptin synthetase B (sypB), and syringopeptin synthetase C
(sypC) genes, complete cds.
ACCESSION AF286216
VERSION AF286216.2 GI:29165621
KEYWORDS
SOURCE Pseudomonas syringae pv. syringae
ORGANISM Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
Pseudomonas syringae pv. syringae
Pseudomonas syringae pv. syringae
Pseudomonadaceae; Pseudomonas.
REFERENCE 1 (bases 1 to 3920)
AUTHORS Scholz-Schroeder,B.K., Hutchison,M.L., Grgurina,I. and Gross,D.C.
TITLE The contribution of syringopeptin and syringomycin to virulence of
Pseudomonas syringae pv. syringae strain B301D on the basis of sypA
and sypB biosynthesis mutant analysis
JOURNAL Mol. Plant Microbe Interact. 14 (3), 336-348 (2001)
MEDLINE 21171030
PUBMED 11277431
REFERENCE 2 (bases 1 to 73800)

Scholz-Schroeder,B.K., Soule,J.D. and Gross,D.C.
TITLE The sypA, sypB, and sypC synthetase genes encode twenty-two modules
involved in the nonribosomal peptide synthesis of syringopeptin by
Pseudomonas syringae pv. syringae B301D
JOURNAL Mol. Plant Microbe Interact. 16, 271-280 (2003)
MEDLINE 93106948
PUBMED 8093238
REFERENCE 3 (bases 1 to 3920)
AUTHORS Scholz-Schroeder,B.K., Soule,J.D., Grgurina,I. and Gross,D.C.
TITLE Direct Submission
JOURNAL Submitted (11-JUN-2000) Department of Plant Pathology, Washington
State University, P.O. Box 646430, Pullman, WA 99163-6430, USA
COMMENT On Mar 24, 2003 this sequence version replaced gi:9802393.
FEATURES
source
1..73800
/organism="Pseudomonas syringae pv. syringae"
/mol_type="genomic DNA"
/strain="B310D"
/db_xref="taxon:321"
/note="pathovar: syringae"
138..16277
/gene="sypA"
138..16277
/gene="sypA"
/function="peptide synthesis"
/codon_start=1
/transl_table=11
/product="syringopeptin synthetase"
/protein_id="AAF99707.2"
/db_xref="GI:29165622"
/translation="MALALTAAGQNIWLDQMTQDGPVLYNIGHLEIDGALDYELFOR
AVDLIAKDALRMVLLDQDEGVPVLAQFATSMPPVPTIDLOGDDPFOAEQWMO
KOLERAFRLDGEPLFRHLKLEQRFYCMYIAHHLALDQWIGIDGLCASLSELYNALH
DOHTPDLTAPSYVDFIKDQSVGSRFALDQTYWLDKRIPEPLVPRPAQION
EAPSRGSLSYGLARDLEQOIIDLALSINASSFHVLLAVLYFTTRAYQRELVIGLP
ILNRPNASYKKTGLFTQLSVRLKFSVDLPFAELVQGIQRAKODYRSORFVSELN
RLELRRAERTQLFDLTFSEYERNHLLRFGQATAHAVYSSNNHEQTPLAHLRTNAYE
AESRLHYIYNEAYFOHDEIAPLARHLRLHBOGLANPHLPVHGSITGMTDTPFKQTV
NAGQPHFANDLTIHQFEARAERPDAAALVEEQTLSYSELNARANQVAHRLIALGV
RPDDRVAICVERGAMIIIGLIGIKSGAGVPLDPAYPLERLATTIGDSAPVALLSQ
SVOPALPVSDVPLIYLDVDLDQDERACNPQSVSPSILAVITYTSGSTGLPKGVNTEH
EDFYNLCSAGTVLNQTPSAFRQILIAAQDKEQVHSLQVIFGGEALETAMLKWYA
RKVNASTQLVNMYGITETTVHTYVPLQPEDARVQASPIGKRIPDLQLYLLDAYGHP
LPPGAVGELVGGAGVARGVRLNDELNTRFLDDPFASAPGARMYRSGDLGRWLDGS
LEYLGRNDEQYKIRGFRIELGEIERLSACEGVRAVAVVREDDEPQRLVAAYVIGTA
GQEPDATWLRQRLSLAEHMLPSAFVSLAEPLITANGLDKRALPVPPTADAPASQY
EABEGLVETRLATLWVELLIGVEQVRHDOFELIGHGSMIAVKLIBRMNDIGLNADVRV
LFGQPTLASLAASGKTTDVQVPASRIPADQORITPNMLPLARLSQAQIDRVAAVAVG
GVSNDQDIYGLVPLQEGILYHHLSSGEGDPYLLQALRLNRSFEQLLDFTALQOVADR
HDILRTAVAWSELDEPQVWVRQARLVEAFLEPHPDQGVAAQLQAFDPRIRIMDIH
QAPMRRLHYAEDPVNQSWAVVLFPHLLDDATSLALGAEIEAFRQGRGNHLPASVRY
RNHVAQARLSMSREEHEAFFRDMQADVDEPTLFFGLQDVQDGSQVAAPTPLFSALMN
GRLRHARRLGVSNASLHHLMAQVGRLSGRQDVFFGYTLGRLLSGQGAERALGME
INTLPLRVAAGQVNEAAVRTTHARLALVSHHEHAPVLSLQGCSSGVAAPTPLFSALMN
YRHVAVGACTPQQAQOTRTWAGVEVLGGEERTNYPLSLSDLDGDSFGLTVQAVAGID
ARRICGYMTVLEQLADALDSRPLASLSDMLPAHERQLLEDENAFDRDPQDLIL
HQLFEAQAAQDPSVAVTYQGGRLSYAELNOMANQIAHRLIAAGVADDRVAICVERS
LEMIAGLVILKAGAGVPLDPSPYERLALMEDIAPKMLLTQGRERFAQASMPV
LLEADARAENGIDRAPVTNPHPAGIGAHLAYLITYSGSTGQPKGVAMPAPLVNLM
HICEQRIERLYMECIALQALAAVVTPEQECCLQDVITAGEQRLITEPMRQFARL
NDARLHNHYGPNESHVATLTALTDGQWTPVLTSPISQAPVANTRIYLTDEHMRPVPVGV
AGELYTGACVARGYLNRLNDLTARETIADPPAASDAGNPRTRLYXTGDLGCMQADGRI
QVPEADPLRLHLQALLPDMYMPAYVHLEKLPVSPNGKDLRALPQPSADAFVSQV
OAPLGDTEETTLAALQCEVLSEVERGODHEFELGSLGSLAVKILERMROVLSADRVY
LFGQPTLAALAAVGGIEIEVPANGIPGCRHITPDMLTITLALDQASIERIVADIPG
GANVQDIYPLAPLQEGILYHHLIAEAGQDPYLLQATFALRDBAHVRRFAEGLQAMIDR

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2004, 11:20:12 ; Search time 341.489 Seconds

(without alignments)
2635.552 Million cell updates/sec

Title: US-10-062-875A-8

Perfect score: 22

Sequence: 1 gtcaagacgctggaagagctc 22

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl: *
1: gb_ba: *
2: gb_htg: *
3: gb_in: *
4: gb_om: *
5: gb_ov: *
6: gb_pat: *
7: gb_ph: *
8: gb_pl: *
9: gb_pr: *
10: gb_ro: *
11: gb_sts: *
12: gb_sy: *
13: gb_un: *
14: gb_vi: *
15: em_ba: *
16: em_fun: *
17: em_hum: *
18: em_in: *
19: em_mu: *
20: em_om: *
21: em_or: *
22: em_ov: *
23: em_pat: *
24: em_ph: *
25: em_pl: *
26: em_ro: *
27: em_sts: *
28: em_un: *
29: em_vi: *
30: em_htg_hum: *
31: em_htg_inv: *
32: em_htg_other: *
33: em_htg_mus: *
34: em_htg_pln: *
35: em_htg_rod: *
36: em_htg_mam: *
37: em_htg_vrt: *
38: em_sy: *
39: em_htgo_hum: *
40: em_htgo_mus: *
41: em_htgo_other: *

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	22	100.0	1306	1	BPI1001	X66858 B.parapertu
2	18.8	85.5	216844	2	AC130788	AC130788 Bos tauru
3	18.8	85.5	302050	1	AL935256	AL935256 Lactobaci
4	18.4	83.6	299910	1	AP005957	AP005957 Bradyrhiz
5	17.8	80.9	522	5	AF125049	AF125049 Oncothyne
6	17.8	80.9	522	5	AF125050	AF125050 Oncothyne
7	17.8	80.9	12946	1	AE013592	AE013592 Methanosa
8	17.8	80.9	95660	2	AC131181	AC131181 Homo sapi
9	17.8	80.9	98914	2	AC114724_3	Continuation (4 of
10	17.8	80.9	110000	2	AC114724_2	Continuation (3 of
11	17.8	80.9	142372	9	AC132152	AC132152 Homo sapi
12	17.8	80.9	151829	2	AC132890	AC132890 Mus muscu
13	17.8	80.9	156817	9	AC009406	AC009406 Homo sapi
14	17.8	80.9	158341	9	AC021733	AC021733 Homo sapi
15	17.8	80.9	158430	2	AC012431	AC012431 Homo sapi
16	17.8	80.9	162577	2	AL954709	AL954709 Danio rer
17	17.8	80.9	186233	9	AC092329	AC092329 Homo sapi
18	17.8	80.9	189317	2	AC024483	AC024483 Homo sapi
19	17.8	80.9	199208	2	AC134427	AC134427 Mus muscu
20	17.8	80.9	252115	2	AC131854	AC131854 Rattus no
21	17.8	80.9	254538	2	AC097032	AC097032 Rattus no
22	17.4	79.1	1625	6	AX399907	AX399907 Sequence
23	17.4	79.1	1653	9	HSACRAP	Z33905 H.sapiens g
24	17.4	79.1	1671	9	AF449218	AF449218 Homo sapi
25	17.4	79.1	11599	1	AE009650	AE009650 Brucella
26	17.4	79.1	13813	1	AE014610	AE014610 Brucella
27	17.4	79.1	79516	9	AC004834	AC004834 Homo sapi
28	17.4	79.1	118106	2	AC110867	AC110867 Rattus no
29	17.4	79.1	170382	2	AC101772	AC101772 Mus muscu
30	17.4	79.1	199658	2	AC133053	AC133053 Rattus no
31	17.4	79.1	210359	9	HSAA430K20	AL121898 Human DNA
32	17.4	79.1	223442	10	AL662804	AL662804 Mouse DNA
33	17.4	79.1	239464	2	AC103079	AC103079 Rattus no
34	17.4	79.1	242506	2	AC094821	AC094821 Rattus no
35	17.2	78.2	2736	10	BC014812	BC014812 Mus muscu
36	17.2	78.2	12531	1	AB011897	AB011897 Xanthomon
37	17.2	78.2	13829	1	AE005935	AE005935 Caulobact
38	17.2	78.2	27818	1	AF030414	AF030414 Gluconace
39	17.2	78.2	64924	2	AC118046	AC118046 Mus muscu
40	17.2	78.2	79676	8	AC090435	AC090435 Chlamydom
41	17.2	78.2	98221	9	AL161940	AL161940 Human DNA
42	17.2	78.2	102653	8	AC090436	AC090436 Chlamydom
43	17.2	78.2	105872	2	AC091993	AC091993 Homo sapi
44	17.2	78.2	110000	2	IMFLCHR34_14	Continuation (15 o
45	17.2	78.2	126394	9	AC109998	AC109998 Homo sapi

ALIGNMENTS

RESULT 1
BPI1001
LOCUS BPI1001 1306 bp DNA linear BCT 07-JUL-2002
DEFINITION B.parapertussis insertion sequence IS1001 tnpA gene for
transposase.
ACCESSION X66858 S51601
VERSION X66858.1 GI:39755
KEYWORDS insertion sequence; insertion sequence IS1001; tnpA gene;
transposase.
SOURCE Bordetella parapertussis
ORGANISM Bordetella parapertussis
Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
Alcaligenaceae; Bordetella.
REFERENCE
AUTHORS 1 van der Zee,A., Agterberg,C., van Agterveld,M., Peeters,M. and

TITLE Mooi, F.R.
Characterization of IS1001, an insertion sequence element of
Bordetella parapertussis
JOURNAL J. Bacteriol. 175 (1), 141-147 (1993)
MEDLINE 93106948
PUBMED 8093238
REFERENCE 2 (bases 1 to 1306)
AUTHORS van der Zee, A.
TITLE Direct Submission
JOURNAL Submitted (12-JUN-1992) A. Van Der Zee, National Institute of Health
&, Environmental Protection, A Van Leeuwenhoeklaan 9, P O Box 1,
3720 BA Bilthoven, THE NETHERLANDS
COMMENT On Jul 10, 2002 this sequence version replaced gi:262329.
FEATURES
source
1. 1306
/organism="Bordetella parapertussis"
/mol_type="genomic DNA"
/db_xref="taxon:519"
/clone="PRPPI"
repeat_region 1. 1306
/insertion_seq="IS1001"
gene 83. 1303
/gene="TnpA"
83. 1303
/gene="TnpA"
CDS
/gene="TnpA"
/codon_start=1
/transl_table=1
/product="transposase"
/protein_id="CAA47326.1"
/db_xref="GI:39756"
/db_xref="SWISS-PROT:Q06126"
/translation="MLDRKLMSLGGWQGYGVREVPEDPGRLLSYLKPTAKVMLC
EQGARCRQVHETTVRRVDLPFEYRVVHVPRRLMCEQCGPRLERLAWIGRYOR
VTDLAQACSQLQSSNQAVARFFELGMHTYKTLDKARLASVREPDMSKIEYLAMD
EFLAHKGRYATVVDPIGRQVLMIGPGRSRTARAFEEQLPPGAQRKAVADMTT
AYELEIOAHSPOAEIVYDLFHVAKYGREVIDRVRVDQANQLRQDRPARRIISRWL
LNRNDNLDRQAVRLDELQANQPLTVYVLRDELKRLWFGYGRPAMARQAMHWYEQ
AEQSGIAALNTFAQRLLKGLHGLIARCRHPLNTSIVEGINNTIKVTKRAYGYRQGY
FPLKIRAFPPGNAR"
BASE COUNT 269 a 371 c 420 g 246 t
ORIGIN
Query Match 100.0%; Score 22; DB 1; Length 1306;
Best Local Similarity 100.0%; Pred. No. 8.8;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GTCAAGACGCTGGACAAGGCTC 22
|||||
Db 482 GTCAAGACGCTGGACAAGGCTC 503
RESULT 2
AC130788
LOCUS
DEFINITION Bos taurus clone RP42-331024, WORKING DRAFT SEQUENCE, 6 ordered
pieces.
AC130788
ACCESSION AC130788
VERSION AC130788.2 GI:22748398
KEYWORDS HTG; HTGS PHASE2; HTGS_DRAFT.
SOURCE
ORGANISM Bos taurus
Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;
Bovidae; Bovinae; Bos.
1 (bases 1 to 216844)
REFERENCE
AUTHORS Akhter, N., Antonellis, A., Ayele, K., Beckstrom-Sternberg, S.M.,
Benjamin, B., Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C.,
Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J.,
Haghighi, P., Han, J., Hansen, N., Ho, S.-L., Idol, J.R., Karlins, E.,
Laric, P., Lee-Lin, S.-Q., Legaspi, R., Maduro, O.L., Maduro, V.B.,
Margulies, E.H., Masiello, C., Maskeri, B., Mastrian, S.D.,
McCluskey, J.C., McDowell, J., Paguirigan, C., Pearson, R.,
Portnoy, M.E., Prasad, A., Reddix-Dugue, N., Schueler, M.G., Sison, C.,

TITLE Stantirip, S., Thomas, J.W., Thomas, P.J., Touchman, J.W., Vogt, J.L.,
Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.
JOURNAL NISC Comparative Sequencing Initiative
REFERENCE 2 (bases 1 to 216844)
AUTHORS Green, E.D.
TITLE Direct Submission
JOURNAL Submitted (14-AUG-2002) NIH Intramural Sequencing Center, 8717
Grovemont Circle, Gaithersburg, MD 20877, USA
3 (bases 1 to 216844)
AUTHORS Green, E.D.
TITLE Direct Submission
JOURNAL Submitted (06-SEP-2002) NIH Intramural Sequencing Center, 8717
Grovemont Circle, Gaithersburg, MD 20877, USA
On Sep 6, 2002 this sequence version replaced gi:22218458.
COMMENT
----- Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc_zoo@nhgri.nih.gov
----- Project Information
Center project name: crz
Center clone name: 331024
The sequence data in this record represents an 'enhanced'
version of a phase 2 submission. Specifically, the indicated
order and orientation of each sequence contig has been
established using one or more of the following: read-pair
data from individual subclones, overlaps with neighboring
clones, alignment with available reference sequence (e.g.,
human), and/or confirmation by PCR testing. In addition,
the sequence assembly is based on at least 8X average
coverage in Q20 bases and has been reviewed to rule out
gross misassemblies, the low-quality ends of sequence
contigs have been trimmed away, and each base is associated
with a Phrap-derived quality score.
----- Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 214028 bases at least Q40
Consensus quality: 215115 bases at least Q30
Consensus quality: 215785 bases at least Q20
Insert size: 208000; agarose-fp
Insert size: 216344; sum-of-contigs
Quality coverage: 8.45x in Q20 bases; agarose-fp
Quality coverage: 8.12x in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 6 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* been provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 53614: contig of 53614 bp in length
* 53615 53714: gap of unknown length
* 53715 132668: contig of 78954 bp in length
* 132669 132768: gap of unknown length
* 132769 150539: contig of 17771 bp in length
* 150540 150639: gap of unknown length
* 150640 184326: contig of 33687 bp in length
* 184327 184426: gap of unknown length
* 184427 187144: contig of 2718 bp in length
* 187145 187244: gap of unknown length
* 187245 216844: contig of 29600 bp in length.
* Location/Qualifiers
1. 216844
/organism="Bos taurus"
/mol_type="genomic DNA"
/db_xref="taxon:9913"

BOLLEN ALEX (BE) ; GODFROID EDMOND (BE) ; UNIV BRUXELLES (BE) ;
FAUCONNIER ALAIN (BE)
FEATURES
Location/Qualifiers

.source
1..951
/organism="Bordetella pertussis"
/mol_type="genomic DNA"
/db_xref="taxon:520"
1..951
/note="unnamed protein product"
/codon_start=1
/transl_table=1
/protein_id="CAC09206.1"
/db_xref="GI:10189668"

CDS
/translation="MNTTKHARLTFLRLMVQQLAHQVCVPEARAYGVTAFTVRK
WLGRLAQOAGLADASSRPTVSPRAIPAKALAVELRRRLTOARIAQALGVASST
VSRVLARAGLSHLADLEPAEPVRYEHQAPGDLHIDIKUGRIQRPGRVTGNRDT
VEGAGWDFVFAIDDHARVAFTDIHPDERFSAVOFLKDAVAYYQRLGVTIORTLTDN
GSAFRSRAFAALCHLGIKARFTPRPQTNKAEFTQSLREWAYHTYONSQHRRA
DAMKSWLHMYMHRPHQIGRAVPISRLNDEYNLLTVHT"

BASE COUNT 174 a 350 c 269 g 158 t
ORIGIN

Query Match 100.0%; Score 18; DB 6; Length 951;
Best Local Similarity 100.0%; Pred. No. 81;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCAGTTCCTCAAGGACGC 18
|||||
DB 582 CCAGTTCCTCAAGGACGC 599

RESULT 2 AF349431 1010 bp DNA linear BCT 11-MAY-2001
LOCUS AF349431
DEFINITION Bordetella holmesii IS481-like insertion sequence, partial
sequence.
ACCESSION AF349431
VERSION AF349431.1 GI:13384215
KEYWORDS
SOURCE
ORGANISM
Bordetella holmesii
Bordetella holmesii
Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
Alcaligenaceae; Bordetella.
REFERENCE
1 (bases 1 to 1010)
Reischl, U., Lehn, N., Sanden, G.N. and Loeffelholz, M.J.
Real-time PCR assay targeting IS481 of Bordetella pertussis and
molecular basis for detecting Bordetella holmesii
J. Clin. Microbiol. 39 (5), 1963-1966 (2001)
JOURNAL
MEDLINE
PUBMED
21225601
11326023
2 (bases 1 to 1010)
Reischl, U., Loeffelholz, M.J. and Sanden, G.N.
Direct Submission
Submitted (14-FEB-2001) Institute of Medical Microbiology and
Hygiene, University of Regensburg, Franz-Josef-Strauss-Allee 11,
Regensburg D-93053, Germany
FEATURES
Location/Qualifiers
1..1010
/organism="Bordetella holmesii"
/mol_type="genomic DNA"
/strain="ATCC 51541"
/db_xref="ATCC:51541"
/db_xref="taxon:35814"
/insertion_seq="IS481-like"

BASE COUNT 185 a 365 c 288 g 170 t 2 others
ORIGIN

Query Match 100.0%; Score 18; DB 1; Length 1010;
Best Local Similarity 100.0%; Pred. No. 81;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCAGTTCCTCAAGGACGC 18
|||||

DB 673 CCAGTTCCTCAAGGACGC 690

RESULT 3
BPEPERRA
LOCUS BPEPERRA
DEFINITION B.pertussis insertion sequence with 28 bp terminal inverted repeats
DNA.
ACCESSION M22031
VERSION M22031.1 GI:144060
KEYWORDS insertion sequence.
SOURCE Bordetella pertussis
ORGANISM Bordetella pertussis
Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
Alcaligenaceae; Bordetella.

REFERENCE
1 (bases 1 to 1053)
McLafferty, M.A., Harcus, D.R. and Hewlett, E.L.
Nucleotide sequence and characterization of a repetitive DNA
element from the genome of Bordetella pertussis with
characteristics of an insertion sequence
J. Gen. Microbiol. 134 (Pt 8), 2297-2306 (1988)
JOURNAL
MEDLINE
PUBMED
89310403
2908119
Original source text: B.pertussis DNA.
FEATURES
Location/Qualifiers
1..1053
/organism="Bordetella pertussis"
/mol_type="genomic DNA"
/db_xref="taxon:520"

repeat_region 1..28
/note="inverted repeat"
repeat_region 1026..1053
/note="inverted repeat"
BASE COUNT 201 a 375 c 296 g 181 t
ORIGIN

Query Match 100.0%; Score 18; DB 1; Length 1053;
Best Local Similarity 100.0%; Pred. No. 81;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCAGTTCCTCAAGGACGC 18
|||||
DB 684 CCAGTTCCTCAAGGACGC 701

RESULT 4
S66929
LOCUS S66929
DEFINITION {inverted repeating element RSBP1} [Bordetella pertussis, Genomic,
1053 nt].
ACCESSION S66929
VERSION S66929.1 GI:440882
KEYWORDS
SOURCE
ORGANISM
Bordetella pertussis
Bordetella pertussis
Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
Alcaligenaceae; Bordetella.

REFERENCE
1 (bases 1 to 1053)
Kirillov, M.I., Shumakov, Iu.L., Nechaeva, E.V., Sinhashina, L.N. and
Karataev, G.I.
Nucleotide sequence and properties of an inverted repeating element
of a Bordetella pertussis chromosome
Genetika 29 (8), 1267-1277 (1993)
JOURNAL
MEDLINE
PUBMED
94010257
8405971
GenBank staff at the National Library of Medicine created this
entry [NCBI gibbon 139922] from the original journal article.
This sequence comes from Fig. 4.

FEATURES
Location/Qualifiers
1..1053
/organism="Bordetella pertussis"
/mol_type="genomic DNA"
/db_xref="taxon:520"

BOULEN ALEX (BE) ; GODFROID EDMOND (BE) ; UNIV BRUXELLES (BE) ;
FAUCONNIER ALAIN (BE)
FEATURES
Source
Location/Qualifiers
1..951

/organism="Bordetella pertussis"
/mol_type="genomic DNA"
/db_xref="taxon:520"
1..951
/note="unnamed protein product"
/codon_start=1
/transl_table=1
/protein_id="CAC09206.1"
/db_xref="GI:10189668"

/translation="MNTKHARLTLRLLEMVQQLIAHQVCPEAARAYGVTAPTVRK
WLGRLAQAGLADASSRPTSPRAIAPAKALAYELRKRRLTQARTALGVSAST
VSRVLARAGLSHLADLEPAEPVRYEHPQGLHIDIKLGRICRPHRVGNRDT
VEGAGWDFVFAIDDHARVAFDTHPDERFSPAVQFLKDAVAYYQRLGVTIQRLTDN
GSAFRSRAPALCHELGIKHFTRPYRQPTNGKAERFIQSALREWAYAHYQNSQHRP
DAMKSWLHHYMHRRPHOGIGRAVPISRLNLDENVNLTVHT"

BASE COUNT 174 a 350 c 269 g 158 t
ORIGIN

Query Match 100.0%; Score 23; DB 6; Length 951;
Best Local Similarity 100.0%; Pred. No. 0.4;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGTTCTGTAGTGTGAGCGTA 23
|||||
Db 815 GAGTTCTGTAGTGTGAGCGTA 793

RESULT 2
AF349431/c 1010 bp DNA linear BCT 11-MAY-2001
LOCUS Bordetella holmesii IS481-like insertion sequence, partial
DEFINITION
ACCESSION AF349431
VERSION AF349431.1 GI:13384215
KEYWORDS
SOURCE
ORGANISM
Bordetella holmesii
Bordetella holmesii
Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
Alcaligenaceae; Bordetella.

REFERENCE
AUTHORS
TITLE
1 (bases 1 to 1010)
Reischl,U., Lehn,N., Sanden,G.N. and Loeffelholz,M.J.
Real-time PCR assay targeting IS481 of Bordetella pertussis and
molecular basis for detecting Bordetella holmesii

JOURNAL
MEDLINE
PUBMED
21225601
11326023
2 (bases 1 to 1010)
Reischl,U., Loeffelholz,M.J. and Sanden,G.N.
Direct Submission
Submitted (14-FEB-2001) Institute of Medical Microbiology and
Hygiene, University of Regensburg, Franz-Josef-Strauss-Allee 11,
Regensburg D-93053, Germany

FEATURES
Source
1..1010
/organism="Bordetella holmesii"
/mol_type="genomic DNA"
/strain="ATCC 51541"
/db_xref="ATCC:51541"
/db_xref="taxon:35814"
/insertion_seq="IS481-like"

BASE COUNT 185 a 365 c 286 g 170 t 2 others
ORIGIN

Query Match 100.0%; Score 23; DB 1; Length 1010;
Best Local Similarity 100.0%; Pred. No. 0.4;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGTTCTGTAGTGTGAGCGTA 23
|||||

Db 906 GAGTTCTGTAGTGTGAGCGTA 884

RESULT 3
BORDERIA/c

LOCUS BPETERRA 1053 bp DNA linear BCT 26-APR-1993
DEFINITION B.pertussis insertion sequence with 28 bp terminal inverted repeats
DNA.

ACCESSION M22031
VERSION M22031.1 GI:144060
KEYWORDS insertion sequence.
SOURCE Bordetella pertussis
ORGANISM Bordetella pertussis
Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
Alcaligenaceae; Bordetella.

REFERENCE
AUTHORS
TITLE
1 (bases 1 to 1053)
McLaifferty,M.A., Harcus,D.R. and Hewlett,E.L.
Nucleotide sequence and characterization of a repetitive DNA
element from the genome of Bordetella pertussis with
characteristics of an insertion sequence

JOURNAL
MEDLINE
PUBMED
89310403
2908119
Original source text: B.pertussis DNA.

COMMENT
FEATURES
Source
1..1053
/organism="Bordetella pertussis"
/mol_type="genomic DNA"
/db_xref="taxon:520"

repeat_region 1..28
/note="inverted repeat"
repeat_region 1026..1053
/note="inverted repeat"

BASE COUNT 201 a 375 c 296 g 181 t
ORIGIN

Query Match 100.0%; Score 23; DB 1; Length 1053;
Best Local Similarity 100.0%; Pred. No. 0.4;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGTTCTGTAGTGTGAGCGTA 23
|||||
Db 917 GAGTTCTGTAGTGTGAGCGTA 895

RESULT 4
S66929/c 1053 bp DNA linear BCT 13-JAN-1994
LOCUS {inverted repeating element RSBp1} [Bordetella pertussis, Genomic,
1053 nt].

ACCESSION S66929
VERSION S66929.1 GI:440882
KEYWORDS
SOURCE
ORGANISM
Bordetella pertussis
Bordetella pertussis
Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
Alcaligenaceae; Bordetella.

REFERENCE
AUTHORS
TITLE
1 (bases 1 to 1053)
Kirillov,M.Iu., Shumakov,Iu.L., Nechaeva,E.V., Sinashina,L.N. and
Karataev,G.I.
Nucleotide sequence and properties of an inverted repeating element
of a Bordetella pertussis chromosome

JOURNAL
MEDLINE
PUBMED
94010257
8405971
GenBank staff at the National Library of Medicine created this
entry [NCBI gibdbeg 139922] from the original journal article.
This sequence comes from Fig. 4.

REMARK
GenBank staff at the National Library of Medicine created this
entry [NCBI gibdbeg 139922] from the original journal article.

FEATURES
Source
1..1053
/organism="Bordetella pertussis"
/mol_type="genomic DNA"
/db_xref="taxon:520"

SEQ ID NO:2

CEQ 10 NOV 03

BOULEN ALEX (BE) ; GODFROID EDMOND (BE) ; UNIV BRUXELLES (BE) ;
FAUCONNIER ALAIN (BE)
Location/Qualifiers

FEATURES

source
1. .951
/organism="Bordetella pertussis"
/mol_type="genomic DNA"
/db_xref="taxon:520"
1. .951
/note="unnamed protein product"

CDS

/codon_start=1
/transl_table=1
/protein_id="CAC09206.1"
/db_xref="GI:10189668"
/translation="MNTKHAFLTLRLRMVQQLIAHQVFEARAYGVTAFTVRK
WLGFLAQOAGLADASSRPTVSPRAIPAKALIVELRRLTOARIQALGVASST
VSRVLAAGLSHLADLEPAEPVRYEHQAPGDLHIDIKLGRIGRPGHRTGNRRT
VEGAGWDFVFAIDHARVAFDIDHDERPSPAVQFLKDAVAYQRLGVTIQRLLTDN
GSAFRSRAPFALCHELGIKTRTPYRPTQNGKARFIQSALREWAYHTYQNSQHRA
DAMKSWLHYNWHRPHOGIGRAVPISRLNDEYNLLTVHT"

BASE COUNT
ORIGIN

174 a 350 c 269 g 158 t

Query Match
Best Local Similarity 100.0%; Score 28; DB 6; Length 951;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CACCGCTTACCGACCTTACCGCCAC 28
DB 712 CACCGCTTACCGACCTTACCGCCAC 739

RESULT 2
AF349431 1010 bp DNA linear BCT 11-MAY-2001
LOCUS
DEFINITION
Bordetella holmesii IS481-like insertion sequence, partial
sequence.

ACCESSION
AF349431
AF349431.1 GI:13384215

KEYWORDS
SOURCE

ORGANISM
Bordetella holmesii

Bordetella holmesii
Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
Alcaligenaceae; Bordetella.

REFERENCE
AUTHORS
TITLE
1 (bases 1 to 1010)
Reischl, U., Lehn, N., Sanden, G.N. and Loeffelholz, M.J.
Real-time PCR assay targeting IS481 of Bordetella pertussis and
molecular basis for detecting Bordetella holmesii

JOURNAL
MEDLINE
21225601
11326023

REFERENCE

AUTHORS

TITLE

JOURNAL

Submitted (14-FEB-2001) Institute of Medical Microbiology and
Hygiene, University of Regensburg, Franz-Josef-Strauss-Allee 11,
Regensburg D-93053, Germany

FEATURES
source

1. .1010
Location/Qualifiers

/organism="Bordetella holmesii"
/mol_type="genomic DNA"

/strain="ATCC 51541"
/db_xref="ATCC:51541"

/db_xref="taxon:35814"
/insertion_seg="IS481-like"

BASE COUNT
ORIGIN
185 a 365 c 288 g 170 t 2 others

Query Match
Best Local Similarity 100.0%; Score 28; DB 1; Length 1010;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CACCGCTTACCGACCTTACCGCCAC 28
|||||

DB 803 CACCGCTTACCGACCTTACCGCCAC 830

RESULT 3

BPTEERRA 1053 bp DNA linear BCT 26-APR-1993
LOCUS
DEFINITION
B.pertussis insertion sequence with 28 bp terminal inverted repeats
DNA.

ACCESSION
M22031.1 GI:144060
VERSION
KEYWORDS
SOURCE
Bordetella pertussis
Bordetella pertussis
ORGANISM
Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
Alcaligenaceae; Bordetella.

REFERENCE
AUTHORS
TITLE
1 (bases 1 to 1053)
McLafferty, M.A., Harcus, D.R. and Hewlett, E.L.
Nucleotide sequence and characterization of a repetitive DNA
element from the genome of Bordetella pertussis with
characteristics of an insertion sequence

JOURNAL
MEDLINE
89310403
2908119
COMMENT
Original source text: B.pertussis DNA.
location/Qualifiers

source

1. .1053

/organism="Bordetella pertussis"
/mol_type="genomic DNA"

/db_xref="taxon:520"

repeat_region
1. .28

/note="inverted repeat"

repeat_region
1026. .1053

/note="inverted repeat"

BASE COUNT
ORIGIN

201 a 375 c 296 g 181 t

Query Match
Best Local Similarity 100.0%; Score 28; DB 1; Length 1053;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CACCGCTTACCGACCTTACCGCCAC 28
DB 814 CACCGCTTACCGACCTTACCGCCAC 841

RESULT 4

LOCUS

566929 1053 bp DNA linear BCT 13-JAN-1994
DEFINITION
{inverted repeating element RSBP1} [Bordetella pertussis, Genomic,
1053 nt].

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Bordetella pertussis
Bordetella pertussis
Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
Alcaligenaceae; Bordetella.

REFERENCE

AUTHORS

TITLE

1 (bases 1 to 1053)
Kirillov, M.Iu., Shumakov, Iu.L., Nechaeva, E.V., Sinishina, L.N. and
Karataev, G.I.
Nucleotide sequence and properties of an inverted repeating element
of a Bordetella pertussis chromosome

JOURNAL
MEDLINE
94010257
8408971

REMARK

GenBank staff at the National Library of Medicine created this
entry [NCBI gisdbq 139922] from the original journal article.
This sequence comes from Fig. 4.

FEATURES
source

1. .1053
/organism="Bordetella pertussis"
/mol_type="genomic DNA"
/db_xref="taxon:520"

BOLLEN ALEX (BE) ; GODFROID EDMOND (BE) ; UNIV BRUXELLES (BE) ;
FAUCONNIER ALAIN (BE)

FEATURES
source

Location/Qualifiers
1. .951

/organism="Bordetella pertussis"
/mol_type="genomic DNA"
/db_xref="taxon:520"

CDS

1. .951

/note="unnamed protein product"

/codon_start=1

/transl_table=11

/protein_id="CAC09206.1"

/db_xref="GI:10189668"

/translation="MNTHKARLTFLRLLEWVOOLIAHQVCPEAARAVGTAPTVRK
MLGRFLAAGQAGLADASSRPVSPRAIAPAKALAVELRKRLTQARIQALGVASAST
VSRVLARAGLSHLADLEPAEPVRYEHOAPGDLHIDIKLGRIGRPHRTGNRDT
VEGAGWDFVFAVIDDHARVAFTDHPDERFPAVOFLKDAVAYQRLGVTIQRLLTDN
GSAFRSRAFAALCHELGIKHRTPRYPOTNGKARFIQSALREMAVYAHYQNSQHRA
DAMKSWLHYNMWRPHQIGRAVPISRLNLDEYNLLTVHT"

BASE COUNT
ORIGIN

174 a 350 c 269 g 158 t

Query Match

100.0%; Score 28; DB 6; Length 951;

Best Local Similarity 100.0%; Pred. No. 0.061;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY

1 GACCAATGGCAAGCCGAACGCTTCATC 28

Db

741 GACCAATGGCAAGCCGAACGCTTCATC 768

RESULT 2

AF349431

LOCUS AF349431 1010 bp DNA linear BCT 11-MAY-2001
DEFINITION Bordetella holmesii IS481-like insertion sequence, partial
sequence.

ACCESSION

AF349431

AF349431.1 GI:13384215

KEYWORDS

SOURCE

Bordetella holmesii
Bordetella holmesii

Alcaligenaceae; Bordetella;
Bacteriia; Proteobacteria; Betaproteobacteria; Burkholderiales;

REFERENCE

AUTHORS

1 (bases 1 to 1010)
Reischl,U., Lehn,N., Sanden,G.N. and Loeffelholz,M.J.

Real-time PCR assay targeting IS481 of Bordetella pertussis and
molecular basis for detecting Bordetella holmesii

J. Clin. Microbiol. 39 (5), 1963-1966 (2001)

JOURNAL

21225601
11326023

PUBMED

2 (bases 1 to 1010)
Reischl,U., Loeffelholz,M.J. and Sanden,G.N.

AUTHORS

Submitted (14-FEB-2001) Institute of Medical Microbiology and
Hygiene, University of Regensburg, Franz-Josef-Strauss-Allee 11,
Regensburg D-93053, Germany

JOURNAL

FEATURES

source

1. .1010

/organism="Bordetella holmesii"

/mol_type="genomic DNA"

/strain="ATCC 51541"

/db_xref="ATCC:51541"

/db_xref="taxon:35814"

/insertion_seq="IS481-like"

BASE COUNT 185 a 365 c 288 g 170 t 2 others

ORIGIN

100.0%; Score 28; DB 1; Length 1010;

Best Local Similarity 100.0%; Pred. No. 0.062;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY

1 GACCAATGGCAAGCCGAACGCTTCATC 28

Db

832 GACCAATGGCAAGCCGAACGCTTCATC 859

RESULT 3
BORDERIA

LOCUS

BPPTERRA

DEFINITION B.pertussis insertion sequence with 28 bp terminal inverted repeats
DNA.

M22031

M22031.1 GI:144060

KEYWORDS

SOURCE

ORGANISM

Bordetella pertussis
Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
Alcaligenaceae; Bordetella.

REFERENCE

AUTHORS

1 (bases 1 to 1053)
McLafferty,M.A., Marcus,D.R. and Hewlett,E.L.

Nucleotide sequence and characterization of a repetitive DNA
element from the genome of Bordetella pertussis with
characteristics of an insertion sequence

J. Gen. Microbiol. 134 (Pt 8), 2297-2306 (1988)

JOURNAL

89310403
2908119

PUBMED

Original source text: B.pertussis DNA.
Location/Qualifiers

FEATURES

source

1. .1053

/organism="Bordetella pertussis"

/mol_type="genomic DNA"

/db_xref="taxon:520"

repeat_region 1. .28
/note="inverted repeat"

repeat_region 1026. .1053
/note="inverted repeat"

BASE COUNT

201 a 375 c 296 g 181 t

ORIGIN

Query Match

100.0%; Score 28; DB 1; Length 1053;

Best Local Similarity 100.0%; Pred. No. 0.062;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY

1 GACCAATGGCAAGCCGAACGCTTCATC 28

Db

843 GACCAATGGCAAGCCGAACGCTTCATC 870

RESULT 4

S66929

LOCUS S66929 1053 bp DNA linear BCT 13-JAN-1994
DEFINITION {inverted repeating element RSBp1} [Bordetella pertussis, Genomic,
1053 nt].

ACCESSION

S66929

VERSION

S66929.1 GI:440882

KEYWORDS

Bordetella pertussis
Bordetella pertussis

SOURCE

Bordetella pertussis
Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;

Alcaligenaceae; Bordetella.

REFERENCE

AUTHORS

1 (bases 1 to 1053)
Kirillov,M.Iu., Shumakov,Iu.L., Nechaeva,E.V., Sinashina,L.N. and
Karataev,G.I.

Nucleotide sequence and properties of an inverted repeating element
of a Bordetella pertussis chromosome

Genetika 29 (8), 1267-1277 (1993)

JOURNAL

94010257
8405971

PUBMED

GenBank staff at the National Library of Medicine created this
entry [NCBI gidsq 139922] from the original journal article.
This sequence comes from Fig. 4.

FEATURES

Location/Qualifiers
1. .1053

/organism="Bordetella pertussis"

/mol_type="genomic DNA"

/db_xref="taxon:520"